

**PERBANDINGAN ANALISIS SENTIMEN TWITTER DENGAN METODE
NAIVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE**

(Studi Kasus : Data Tanggapan Tweet Terhadap Mobil Listrik)

SKRIPSI

Untuk Memenuhi Sebagian Persyaratan Mencapai Derajat Sarjana S-1
Program Studi Teknik Informatika



Disusun Oleh:

Nabila Karima Azka

19106050007

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA**

2023



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI

Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-254/Un.02/DST/PP.00.9/01/2023

Tugas Akhir dengan judul : Perbandingan Analisis Sentimen dengan Metode Naive Bayes Classifier dan Support Vector Machine (Studi Kasus : Data Tanggapan Twitter Terhadap Mobil Listrik)

yang dipersiapkan dan disusun oleh:

Nama : NABILA KARIMA AZKA
Nomor Induk Mahasiswa : 19106050007
Telah diujikan pada : Selasa, 24 Januari 2023
Nilai ujian Tugas Akhir : A

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Muhammad Didik Rohmad Wahyudi, S.T., MT.
SIGNED

Valid ID: 63d1fa9eb45d0



Penguji I

Ir. Maria Ulfah Siregar, S.Kom., MIT., Ph.D.
SIGNED

Valid ID: 63d1e7ab39ddf



Penguji II

Dwi Otik Kurniawati, M.Eng.
SIGNED

Valid ID: 63d1f9209564e



Yogyakarta, 24 Januari 2023
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 63d21e9985605

SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi

Lamp : -

Kepada

Yth. Dekan Fakultas Sains dan
Teknologi UIN Sunan Kalijaga
Yogyakarta di Yogyakarta
Assalamualaikum Wr. Wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi
serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing
berpendapat bahwa skripsi Saudara:

Nama : Nabila Karima Azka

NIM : 19106050007

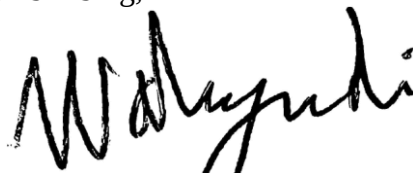
Judul Skripsi : Perbandingan Analisis Sentimen dengan Metode
Naïve Bayes Classifier dan *Support Vector
Machine* (Studi
Kasus : Data Tanggapan Tweet Terhadap Mobil
Listrik)

sudah dapat diajukan kembali kepada Program Studi Teknik Informatika
Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah
satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi
Teknik Informatika.

Dengan ini kami mengharap agar skripsi/tugas akhir Saudara dapat
segera di-*munaqasyah*-kan. Atas perhatiannya kami ucapkan terima kasih.
Wassalamualaikum Wr. Wb.

Yogyakarta, 12 Januari 2023

Pembimbing,



Muhammad Didi Rohmad Wahyudi, S.T., MT.
NIP. 19760812 200901 1015

PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan dibawah ini:

Nama : Nabila Karima Azka
NIM : 19106050007
Program Studi : Teknik Informatika
Fakultas : Sains dan Teknologi

menyatakan bahwa skripsi saya yang berjudul "Perbandingan Perbandingan Analisis Sentimen dengan Metode *Naive Bayes Classifier* dan *Support Vector Machine* (Studi Kasus : Data Tanggapan Tweet Terhadap Mobil Listrik)" merupakan hasil penelitian saya sendiri, tidak terdapat pada karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu perguruan tinggi, dan bukan plagiasi karya orang lain kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 16 Januari 2023

Yang menyatakan,



Nabila Karima Azka

NIM. 19106050007

KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Allah Swt. atas ridhonya saya dapat menyelesaikan penyusunan skripsi ini. Adapun judul skripsi yang saya ajukan adalah “Analisis Sentimen Twitter dengan Metode *Naive Bayes* dan *Support Vector Machine* (Studi Kasus : Data Tanggapan Tweet Terhadap Mobil Listrik)” dapat terselesaikan dengan baik. Tidak dapat disangkal bahwa butuh usaha yang keras dalam penyelesaian pengerjaan skripsi ini. Namun, skripsi ini tidak akan selesai tanpa orang-orang tercinta di sekeliling saya yang telah membantu dan mendukung. Terima kasih saya sampaikan kepada :

1. Bapak Prof. Dr.Phil. Al Makin, S.Ag., M.A., selaku Rektor UIN Sunan Kalijaga Yogyakarta.
2. Ibu Dr. Dra. Hj. Khurul Wardati, M.Si., selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
3. Ibu Ir.Maria Ulfah Siregar, S.Kom., MIT., Ph.D., selaku Ketua Program Studi S1 Teknik Informatika.
4. Bapak Muhammad Didik Rohmad Wahyudi, S.T., MT., selaku Dosen Pembimbing yang telah sabar membimbing, mengarahkan, dan memberi nasihat selama penyusunan skripsi.
5. Kedua Orang tua saya Bapak Muhammad Munawwir, Ibu Apriyatun, S.E. dan keluarga yang telah memberikan doa dan dukungan selama proses pembuatan skripsi.
6. Bapak-Ibu Dosen Teknik Informatika UIN Sunan Kalijaga yang telah bersedia untuk membantu dalam memvalidasi data.
7. Seluruh staf dan karyawan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
8. Staf dan karyawan Perpustakaan UIN Sunan Kalijaga Yogyakarta yang telah membantu dalam pencarian referensi penulisan skripsi.
9. Teman-teman seperjuangan Teknik Informatika 2019 UIN Sunan Kalijaga Yogyakarta.

10. Teman-teman yang selalu peduli, saling mendukung selama proses penulisan skripsi Wenny, Hurun'in, Muna, Nazla, Mualim, Museiri
11. Teman diskusi dan seperjuangan skripsi Bahaudin, Nazla, Enny, Kak Esther
12. Teman-teman online yang selalu saling support.
13. Teman-teman dipondok Wahid Hasyim
14. Teman-teman KKN 108 Pujon Kidul
15. Semua pihak yang tidak dapat disebutkan satu persatu yang telah membantu memberi dukungan

Penulis menyadari masih banyak kekurangan dalam skripsi ini, sehingga kritik dan saran yang membangun dari pembaca sangat penulis harapkan. Semoga skripsi ini memberikan manfaat untuk mendorong penelitian-penelitian selanjutnya.

Yogyakarta, Januari 2023

Penyusun

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Nabila Karima Azka

19106050007

HALAMAN PERSEMBAHAN

Skripsi ini penulis persembahkan untuk almamater, kedua orang tua, keluarga, dan teman-teman seperjuangan yang selalu memberikan semangat support dan semangat kepada penulis serta semua pihak yang turut andil dalam proses penulisan ini yang tidak dapat disebutkan satu persatu.

***Terimakasih
Semuanya.***



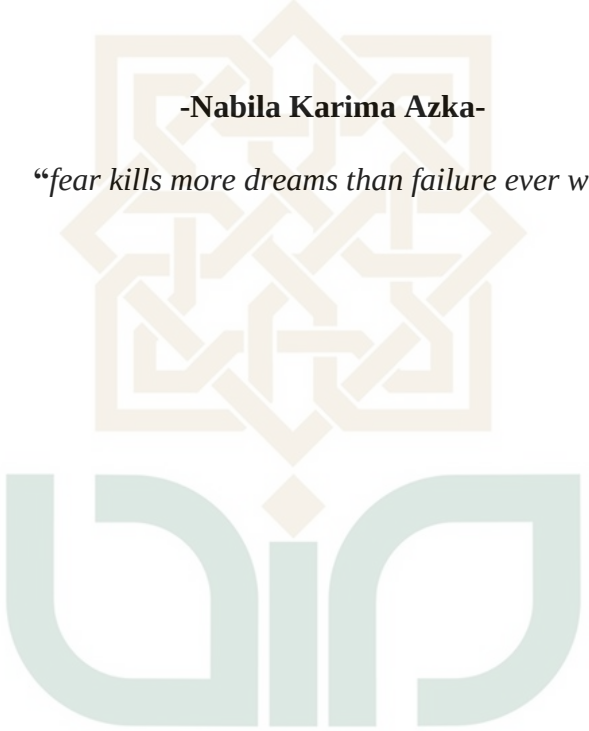
STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN MOTTO

*“You don’t need to be great to start, but you need
to start to be a great”*

-Nabila Karima Azka-

“fear kills more dreams than failure ever will”



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR ISI

PENGESAHAN TUGAS AKHIR	i
SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR	ii
PERNYATAAN KEASLIAN SKRIPSI.....	iii
KATA PENGANTAR	iv
HALAMAN PERSEMBAHAN	vi
HALAMAN MOTTO	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
INTISARI.....	xii
ABSTRACT.....	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Kebaruan Penelitian.....	5
1.7 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	7
2.1 Tinjauan Pustaka	7
2.2 Landasan Teori	12
2.2.1 Mobil Listrik.....	12
2.2.2 Klasifikasi	12
2.2.3 Analisis Sentiman	12
2.2.4 Twitter.....	13
2.2.5 <i>Text Mining</i>	13
2.2.6 Algoritma Naive Bayes.....	14
2.2.7 Support Vector Machine	14

2.2.8 Confusion Matrix.....	16
BAB III METODE PENELITIAN.....	18
3.1 Metode Penelitian	18
3.2 Tahapan Penelitian	18
3.2.1 Studi Pustaka.....	18
3.2.2 Pengumpulan Data	19
3.2.3 Seleksi Data.....	19
3.2.4 Pembersihan Data (Data Preprocessing)	19
3.2.5 Analisis dan Evaluasi	20
3.2.6 Implementasi	21
BAB IV HASIL DAN PEMBAHASAN	22
4.1 Pengumpulan Data.....	22
4.2 Seleksi dan Pelabelan Data.....	25
4.3 Pembersihan Data (<i>Preprocessing Data</i>).....	26
4.3.1 Casefolding dan Cleansing.....	26
4.3.2 Tokenisasi	28
4.3.3 <i>Stopword Removal</i>	30
4.3.4 Stemming	32
4.3.5 Normalisasi	34
4.4 Analisa dan Evaluasi	35
4.4.1 Analisa	35
4.4.2 Evaluasi Model Klasifikasi.....	56
4.5 Implementasi dan Hasil	60
BAB V KESIMPULAN DAN SARAN.....	66
5.1 Kesimpulan.....	66
5.2 Saran	67
DAFTAR PUSTAKA	68

DAFTAR TABEL

Tabel 2. 1 Ringkasan Tinjauan Pustaka	9
Tabel 2. 2 Tabel Confusion Matrix	16
Tabel 4. 1 Contoh Data Komentar Twitter.....	23
Tabel 4. 2 Contoh Data Twitter Yang Sudah Diberi Label.....	25
Tabel 4. 3 Contoh penerapan cleansing dan casefolding	27
Tabel 4. 4 Contoh penerapan tokenisasi	29
Tabel 4. 5 Contoh penerapan stopword removal	30
Tabel 4. 6 Contoh penerapan proses stemming	32
Tabel 4. 7 Contoh penerapan normalisasi	34
Tabel 4. 8 Contoh data twitter pada proses Term Weighting	36
Tabel 4. 9 Contoh pembobotan kata	38
Tabel 4. 10 Nilai TF-IDF	40
Tabel 4. 11 Probabilitas term positif	43
Tabel 4. 12 Contoh perhitungan term negatif	45
Tabel 4. 13 Contoh perhitungan term netral	47
Tabel 4. 14 Contoh Klasifikasi 3 kelas dengan OVR	51
Tabel 4. 15 Contoh data latih One Versus Rest	52
Tabel 4. 16 Nilai x yang didapat dari TF-IDF	53
Tabel 4. 17 Contoh dari nilai Y.....	55
Tabel 4. 18 Jumlah Data Setiap Kelas Sentimen	56

DAFTAR GAMBAR

Gambar 2. 1 Hyperplane terbaik memisahkan kedua class Negatif dan Positif ...	15
Gambar 3. 1 Tahapan Penelitian Sentimen Analisis	18
Gambar 4. 1 Diagram Tahapan Sentimen Analisis	22
Gambar 4. 2 Confusion Matrix Naive Bayes Classifier.....	57
Gambar 4. 3 Confusion Matrix Support Vector Machine.....	58
Gambar 4. 4 Diagram Perbandingan Akurasi Model <i>Naive Bayes Classifier</i>	59
Gambar 4. 5 Hasil Implementasi Naive Bayes Classifier	61
Gambar 4. 6 Hasil Implementasi Model Support Vector Machine.....	61
Gambar 4. 7 Word Cloud Sentimen Positif Klasifikasi Naive Bayes.....	62
Gambar 4. 8 Word Cloud Sentimen Negatif Klasifikasi Naive Bayes.....	62
Gambar 4. 9 Word Cloud Sentimen Netral Klasifikasi Naive Bayes	63
Gambar 4. 10 Word Cloud Sentimen Positif Klasifikasi SVM	63
Gambar 4. 11 Word Cloud Sentimen Negatif Klasifikasi SVM.....	64
Gambar 4. 12 Word Cloud Sentimen Netral Klasifikasi SVM.....	64

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PERBANDINGAN ANALISIS SENTIMEN TWITTER DENGAN METODE NAIVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE

(Studi Kasus : Data Tanggapan Tweet Terhadap Mobil Listrik)

Nabila Karima Azka
19106050007

INTISARI

Perubahan iklim merupakan suatu isu yang menjadi perhatian serius bagi dunia terutama pemerintah Indonesia. Salah satu upaya menekan efek gas rumah kaca adalah mengembangkan alat transportasi yang lebih ramah lingkungan yaitu mobil listrik. Namun banyak pro dan kontra yang diperbincangkan terhadap keefektifan mobil listrik sebagai solusi kendaraan yang lebih ramah lingkungan. Maka dari itu penelitian ini bertujuan untuk menganalisis opini masyarakat atau sentimen yang berkembang di masyarakat terhadap kebijakan mobil listrik menggunakan metode *Naïve Bayes Classifier* dan *Support Vector Machine* pada media sosial twitter.

Tujuan dari dilakukannya perbandingan kedua metode tersebut adalah mengetahui model mana yang memiliki kinerja paling baik. Dipilihnya metode *Naïve Bayes Classifier* dan *Support Vector Machine* pada penelitian ini dikarenakan *Naïve Bayes* merupakan salah satu metode klasifikasi yang paling baik sehingga sering digunakan. Sementara itu *Support Vector Machine* tergolong metode yang masih baru sehingga dinilai lebih baik.

Hasil implementasi pada data uji sebanyak 6.316 data tweet menggunakan model *Naïve Bayes Classifier* menghasilkan sentimen positif sebanyak 54,8% dengan 3859 data, kelas sentimen netral sebanyak 11,4% dengan 805 data, dan kelas sentimen negatif sebanyak 33,8% dengan 2333 data sedangkan menggunakan model *Support Vector Machine* menghasilkan kelas sentimen positif sebanyak 51,7% dengan 3641 data, kelas sentimen netral sebanyak 12,4% dengan 876 data, dan kelas sentimen negatif sebanyak 35,9% dengan 2530 data. Total data yang digunakan pada penelitian ini berjumlah 13.364 data tweet dari twitter, yang mana 6.316 digunakan sebagai data latih dengan label dan 7.048 data uji tanpa label. Akurasi yang didapatkan pada model *Naïve Bayes Classifier* adalah 73,18% sedangkan pada model *Support Vector Machine* akurasinya adalah 80,85%.

Kata kunci : Mobil listrik, Analisis Sentimen, *Naïve Bayes Classifier*, *Support Vector Machine*

COMPARISON OF TWITTER SENTIMENT ANALYSIS WITH NAIVE BAYES CLASSIFIER METHOD AND SUPPORT VECTOR MACHINE

(Case Study: Data on Tweet Responses to Electric Cars)

Nabila Karima Azka

19106050007

ABSTRACT

Climate change is an issue that is of serious concern to the world, especially the Indonesian government. One of the efforts to suppress the greenhouse gas effect is to develop a more environmentally friendly mode of transportation, namely an electric car. However, many pros and cons have been discussed about the effectiveness of electric cars as a more environmentally friendly vehicle solution. Therefore this study aims to analyze public opinion or sentiments that develop in society towards electric car policies using the *Naïve Bayes Classifier* and *Support Vector Machine* methods on Twitter social media.

The purpose of doing a comparison of the two methods is to find out which model has the best performance. The *Naïve Bayes Classifier* and *Support Vector Machine* methods were chosen in this study because *Naïve Bayes* is one of the best classification methods so it is often used. Meanwhile, the *Support Vector Machine* is a relatively new method, so it is considered better.

The results of the implementation of the test data of 6,316 tweet data using the *Naïve Bayes Classifier* model produce positive sentiments of 54.8% with 3859 data, 11.4% neutral sentiment class with 805 data, and 33.8% negative sentiment class with 2383 data while using the *Support Vector Machine* model it produces a positive sentiment class of 51.7% with 3641 data, a neutral sentiment class of 12.4% with 876 data, and a negative sentiment class of 35.9% with 2530 data. The total data used in this study amounted to 13,364 data tweets from Twitter, of which 6,316 were used as training data with labels and 7,048 test data without labels. The accuracy obtained in the *Naïve Bayes Classifier* model is 73.18% while in the *Support Vector Machine* model the accuracy is 80.85%.

Keywords : Electric car, Sentiment Analysis, *Naïve Bayes Classifier*, *Support Vector Machine*

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perubahan iklim merupakan salah satu isu yang sering dibahas oleh berbagai negara akhir-akhir ini. Peningkatan suhu bumi, kekeringan, meningkatnya volume laut, punahnya beberapa spesies, merupakan beberapa akibat yang disebabkan oleh perubahan iklim. Salah satu cara terbaik untuk menekan ekstremnya perubahan iklim adalah dengan mengurangi emisi gas rumah kaca. Jika perubahan iklim di seluruh dunia terus meningkat dan potensi banyak bencana yang akan terjadi maka mengakibatkan kerugian yang sangat besar bagi peradaban manusia. Oleh sebab itu, negara di berbagai dunia harus tegas mengambil kebijakan untuk mendukung kegiatan yang ramah lingkungan demi menekan perubahan iklim di dunia.

Tanpa kita sadari, gas rumah kaca merupakan penyumbang terbesar dalam hal kenaikan suhu rata-rata di bumi akibat kegiatan yang dilakukan manusia. Sebagai contoh, penggunaan kendaraan pribadi untuk transportasi masih menggunakan bahan bakar fosil. Energi fosil merupakan energi yang tak terbarukan dan akan habis dalam beberapa tahun kedepan. Maka dari itu, diperlukannya pembaharuan dalam penggunaan sumber energi dunia. Mobil listrik merupakan salah satu sarana transportasi yang bisa memenuhi mobilitas masyarakat namun tetap ramah lingkungan karena tidak memiliki polusi atau emisi gas buang (Aziz dkk.). Mobil listrik mempunyai kelebihan yaitu dapat mengurangi emisi gas rumah kaca karena tidak memiliki gas buang.

Kehadiran mobil listrik menimbulkan pro dan kontra yang ramai diperbincangkan di media sosial twitter. Twitter merupakan media sosial yang hadir dengan format yang berbeda, di mana twitter memiliki konsep yaitu menyebarkan informasi pesan secara singkat, padat dan real time dengan kalimat yang kurang dari 140 karakter kepada pembacanya di seluruh dunia yang bisa digunakan sebagai sarana penyebar informasi kepada semua orang

baik yang dikenal maupun tidak, untuk memberitahukan keberadaan penggunaannya(Annisa Batu Bara dkk.).

Dengan menggunakan Twitter, seluruh orang dapat terhubung dengan teman dan keluarga dengan menggunakan komputer atau ponsel mereka. Media sosial ini pertama kali didirikan pada Maret 2006 dan dikembangkan oleh Jack Dorsey, seorang mahasiswa sarjana asal New York. Sebagai salah satu media sosial yang paling populer di Indonesia, twitter memiliki jumlah pengguna yang cukup banyak di Indonesia. Tercatat pada bulan Januari 2022, Indonesia menempati jumlah pengguna twitter terbanyak kelima di dunia dengan 18,45 juta pengguna. Dilihat dari data tersebut, dapat disimpulkan bahwa twitter merupakan media sosial yang memiliki potensi besar bagi pengguna di Indonesia. Melalui twitter, pengguna dapat mengemukakan opini, berbagi kegiatan sehari-hari, berbagi foto, serta dapat menggunakan *space* jika ingin berbincang langsung dengan pengguna lainnya.

Biasanya pengguna twitter ramai menanggapi topik yang sedang hangat dibahas. Jika banyak pengguna yang menggunakan suatu kata kunci tersebut, maka akan menjadi *trending topic* di twitter yang dapat berubah secara *realtime*. Banyaknya pengguna twitter yang menyampaikan opini mereka di twitter menjadi salah satu media untuk menganalisis kecenderungan opini masyarakat terhadap suatu topik yang sedang hangat dibicarakan. Proses analisis opini masyarakat di twitter guna memperoleh kecenderungan opini masyarakat terhadap suatu isu disebut analisis sentimen. Nantinya, opini masyarakat dapat dikelompokkan menjadi opini positif, negatif, atau netral.

Analisis sentimen merupakan sebuah metode yang digunakan untuk mengekstrak data opini, memahami serta mengolah tekstual data secara otomatis untuk melihat sentimen yang terkandung dalam sebuah opini(Sari & Wibowo, 2019). Sentimen analisis sering dilakukan pada data tekstual untuk mengetahui sentimen atau respon masyarakat terhadap suatu topik atau produk. Salah satu analisis sentimen yang bisa dilakukan adalah analisis sentimen pada isu yang terjadi pada media sosial twitter. Sebagai salah satu

isu yang sering dibahas di twitter karena menimbulkan pro dan kontra, maka diperlukan adanya analisis sentimen terhadap hadirnya mobil listrik sebagai gebrakan pemerintah dalam upaya menekan perubahan iklim dunia.

Pada penelitian sebelumnya yang dilakukan oleh Ahmad Santoso, Agung Nugroho dan Aswan Supriyadi Sunge yang berjudul “Analisis Sentimen Tentang Mobil Listrik Dengan Metode *Support Vector Machine* Dan *Feature Selection Particle Swarm Optimization*”(Santoso dkk.). Pada penelitian ini membahas tentang analisis sentimen terhadap mobil listrik menggunakan metode *Support Vector Machine* Dan *Feature Selection Particle Swarm Optimization*. Hasil dari penelitian tersebut adalah dapat diketahui bahwa 94,25% pengguna twitter setuju dan 5,75% pengguna twitter tidak setuju terhadap kehadiran mobil listrik. Penggunaan *feature selection Particle Swarm Optimization* pada metode *Support Vector Machine* untuk menganalisis sentimen masyarakat mengenai mobil listrik dapat meningkatkan nilai akurasi dan AUC.

Berdasarkan hasil penelitian terdahulu, penulis tertarik untuk melakukan perbandingan menggunakan metode *Naive Bayes* dan *Support Vector Machine* yang akan penulis gunakan dalam analisis sentimen data twitter dikarenakan *Naive Bayes* memiliki kelebihan yaitu sederhana, cepat dan memiliki akurasi yang tinggi. Sedangkan *Support Vector Machine* mampu mengidentifikasi *hyperplane* terpisah yang memaksimalkan margin antara dua kelas yang berbeda. Tujuan dilakukan perbandingan dua metode ini adalah untuk mengetahui metode mana yang lebih akurat digunakan untuk analisis sentimen twitter.

1.2 Rumusan Masalah

Berdasarkan penjelasan latar belakang diatas, maka rumusan masalah yang akan dibahas adalah :

1. Bagaimana mengimplementasikan algoritma *Naive Bayes* dan *Support Vector Machine* pada analisis sentimen data twitter?
2. Bagaimana kecenderungan opini masyarakat terhadap mobil listrik?

3. Bagaimana perbandingan kedua algoritma *Naive Bayes* dan *Support Vector Machine* dalam analisis sentimen terhadap mobil listrik?

1.3 Batasan Masalah

Batasan-batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Data tweet yang digunakan adalah data tweet bahasa Indonesia dengan kata kunci “Mobil listrik”.
2. Data yang diambil adalah data random tahun 2022 sebanyak 15.000 *tweet* pada periode Agustus-Oktober.
3. Bahasa asing, bahasa daerah dan emoji pada tweet diabaikan.
4. Menggunakan bahasa pemrograman Python.
5. Penelitian ini melakukan klasifikasi sentimen menjadi tiga kelas yaitu positif, negatif dan netral.
6. Algoritma yang digunakan dalam pengklasifikasian ini adalah *Naive Bayes Classifier* dan *Support Vector Machine*.
7. Pengujian validasi data dilakukan menggunakan metode *cross validation*.
8. Pengukuran akurasi dilakukan menggunakan *confusion matrix*.

1.4 Tujuan Penelitian

Adapun tujuan dari dilakukannya penelitian tugas akhir ini adalah:

1. Mengetahui sejauh mana tingkat akurasi analisis sentimen menggunakan metode *Naive Bayes Classifier* dan *Support Vector Machine*.
2. Mengklasifikasikan sentimen mengenai tanggapan terhadap mobil listrik ke dalam tiga kelas yaitu positif, negatif, dan netral.

1.5 Manfaat Penelitian

Dengan tercapainya tujuan penelitian di atas, maka diharapkan dapat bermanfaat bagi semua pihak, manfaat dari penelitian ini adalah:

1. Dengan menggunakan dua metode yang berbeda dapat diketahui sentimen masyarakat terhadap mobil listrik serta dapat mengetahui perbedaan kedua metode tersebut dari sentimen analis yang dihasilkan.
2. Bagi UIN Sunan Kalijaga, dapat digunakan sebagai bahan referensi yang terkait dengan peningkatan pengguna media sosial seiring dengan berkembangnya teknologi informasi.
3. Bagi Penulis, dapat memberi pengetahuan yang baru, bahwa dari data tweet bisa dijadikan sumber data untuk mengetahui opini publik terhadap suatu isu.
4. Bagi Pembaca, dengan adanya penulisan ini diharapkan bermanfaat sebagai sumber pengetahuan dan untuk dijadikan referensi pada penelitian yang selanjutnya akan dikembangkan.

1.6 Kebaruan Penelitian

Penelitian mengenai sentimen analisis mobil listrik sudah pernah dilakukan sebelumnya. Namun belum ada yang meneliti dengan menggunakan perbandingan metode *Naïve Bayes Classifier* dan *Support Vector Machine*.

1.7 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam tugas akhir ini terbagi dalam beberapa pokok bahasan, yaitu:

BAB I PENDAHULUAN

Bab ini membahas latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan tugas akhir.

BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI

Bab ini menjelaskan penelitian yang telah dilakukan sebelumnya serta landasan teori yang berhubungan dengan topik yang akan dibahas pada penelitian ini.

BAB III METODE PENELITIAN

Bab ini membahas metode yang digunakan dalam penelitian serta langkah-langkah yang harus dilakukan untuk mencapai tujuan dalam menyelesaikan penelitian tugas akhir.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan proses yang dilakukan dalam penelitian mulai dari pengumpulan data hingga hasil yang didapatkan dari penelitian.

BAB V KESIMPULAN DAN SARAN

Bagian ini berisi kesimpulan yang didapatkan dari penelitian tugas akhir dan saran yang dapat diterapkan untuk pembangunan lebih lanjut nantinya.



BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari penelitian diatas dapat disimpulkan bahwa data yang digunakan merupakan data yang diambil dari *Twitter* dengan cara scraping data menggunakan library *pyhton snsrape* dengan kata kunci “mobil listrik”. Data yang didapatkan sebanyak 15.000 data tweet yang kemudian dilakukan *preprocessing data* sehingga menyisakan 13.364 data. Dari 13.364 data ini kemudian dilakukan seleksi data sehingga mendapatkan 6316 data latih dan 7.048 uji. Data diolah menggunakan model *Naïve Bayes Classifier* dan *Support Vector Machine* dengan pembobotan TF-IDF serta dilakukan evaluasi model dengan menggunakan *confusion matrix* dan *cross validation*.

Berdasarkan penelitian yang telah dilakukan terhadap 6.316 data latih dengan pembobotan TF-IDF menggunakan model *Naïve Bayes Classifier* didapatkan nilai rata-rata akurasi saat menggunakan *cross validation* sebesar 73,05%. Sedangkan nilai rata-rata akurasi saat menggunakan *cross validation* pada model *Support Vector Machine* adalah sebesar 80,38%.

Dari perbandingan hasil akurasi diatas dapat disimpulkan bahwa kombinasi penggunaan metode untuk mencari model klasifikasi menggunakan *Naïve Bayes Classifier* dan *Support Vector Machine*, metode SVM lah yang memiliki akurasi yang lebih tinggi dibandingkan dengan metode NBC.

Selanjutnya, mengimplementasikan hasil tersebut pada 7.048 data uji yang belum memiliki label. Untuk model *Naïve Bayes Classifier* sentimen positif sebanyak 54,8% dengan 3859 data, kelas sentimen netral sebanyak 11,4% dengan 805 data, dan kelas sentimen negatif sebanyak 33,8% dengan 2383 data. Sedangkan untuk model *Support Vector Machine* dengan kernel RBF menghasilkan kelas sentimen positif sebanyak 51,7% dengan 3641 data, kelas sentimen netral sebanyak 12,4% dengan 876 data, dan kelas sentimen

negatif sebanyak 35,9% dengan 2530 data. Dari kedua model tersebut sama-sama menjadikan kelas sentimen positif sebagai nilai tertinggi.

Dengan adanya hasil sentimen diatas, kita dapat menyimpulkan bahwa tanggapan masyarakat terhadap adanya kebijakan mobil listrik adalah suatu hal yang positif. Diharapkan dengan adanya hasil penelitian ini masyarakat maupun pemerintah dapat mendukung upaya dalam pengurangan pencemaran udara dengan mulai menggunakan mobil listrik.

5.2 Saran

Pada penelitian ini tentunya tidak terlepas dari kelemahan dan kekurangan. Oleh karena itu peneliti menyarankan adanya perbaikan-perbaikan yang dapat dijadikan masukan untuk peneliti selanjut, adapun sebagai berikut :

1. Penelitian selanjutnya diharapkan menggunakan data latih dengan jumlah yang lebih banyak agar mendapatkan model yang lebih akurat.
2. Penelitian selanjutnya diharapkan bisa menggunakan tahapan pembersihan data yang berbeda.
3. Penelitian selanjutnya diharapkan dapat melakukan proses labeling data dengan menggunakan dua atau lebih personil agar dapat memberi sudut pandang yang berbeda.
4. Penelitian selanjutnya diharapkan menggunakan pengembangan dari algoritma *Naïve Bayes Classifier*.

DAFTAR PUSTAKA

- Amelia, R., Sesty Prastiwi, N., Eka Purbaya, M., & informatika, F. (2022). Impementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Masyarakat Indonesia Mengenai Drama Korea Pada Twitter. *Jurnal Riset Komputer*, 9(2), 2407–389. <https://doi.org/10.30865/jurikom.v9i2.3895>
- Annisa Batu Bara, E., Amelia Nasution, K., Zahara Ginting, R., & Studi Ilmu Perpustakaan, P. (t.t.). *Penelitian tentang Twitter*.
- Ashari Muin, A. (2016). Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi). *Jurnal Ilmiah Ilmu Komputer*, 2(1). <http://ejournal.fikom-unasman.ac.id>
- Aziz, M., Marcellino, Y., Agnita Rizki, I., Anwar Ikhwanuddin, S., & Welman Simatupang, J. (t.t.). *STUDI ANALISIS PERKEMBANGAN TEKNOLOGI DAN DUKUNGAN PEMERINTAH INDONESIA TERKAIT MOBIL LISTRIK* (Vol. 22, Issue 1).
- Devita, R. N., Herwanto, H. W., & Wibawa, A. P. (2018). Perbandingan Kinerja Metode Naive Bayes dan K-Nearest Neighbor untuk Klasifikasi Artikel Berbahasa indonesia. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 5(4), 427. <https://doi.org/10.25126/jtiik.201854773>
- Farid Rifai, M., Jatnika, H., Valentino, B., & Tinggi Teknik PLN, S. (2019). *Penerapan Algoritma Naïve Bayes Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi Microsoft Office Specialist (MOS)*. 12(2).
- Fikri, M. I., Sabrila, T. S., Azhar, Y., & Malang, U. M. (t.t.). *Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter*.

- Homepage, J., Adha, T., Putri, M., Enri, U., & Sari, B. N. (2020). IJCIT (Indonesian Journal on Computer and Information Technology) Analisis Algoritma Naive Bayes Classifier untuk Klasifikasi Tweet Pelecehan Seksual dengan #MeToo. Dalam *IJCIT (Indonesian Journal on Computer and Information Technology)* (Vol. 5, Issue 2).
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. Dalam *Jurnal Sains Komputer & Informatika (J-SAKTI)* (Vol. 5, Issue 2).
- Nugroho, A. S. (2008). SUPPORT VECTOR MACHINE: PARADIGMA BARU DALAM SOFTCOMPUTING. Dalam *Konferensi Nasional Sistem dan Informatika*.
- Nurjannah, M., Fitri Astuti, I., & Program Studi, D. (2013). PENERAPAN ALGORITMA TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY (TF-IDF) UNTUK TEXT MINING Mahasiswa S1 Program Studi Ilmu Komputer FMIPA Universitas Mulawarman 2,3). Dalam *Jurnal Informatika Mulawarman* (Vol. 8, Issue 3).
- Santoso, A., Nugroho, A., & Sunge, A. S. (t.t.). Analisis Sentimen Tentang Mobil Listrik Dengan Metode Support Vector Machine Dan Feature Selection Particle Swarm Optimization. Dalam *Journal of Practical Computer Science* (Vol. 2, Issue 1).
- Sari, F. V., & Wibowo, A. (2019). ANALISIS SENTIMEN PELANGGAN TOKO ONLINE JD.ID MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER BERBASIS KONVERSI IKON EMOSI. *Jurnal SIMETRIS*, 10(2).

Sihombing, P. R., & Yuliati, I. F. (2021). Penerapan Metode Machine Learning dalam Klasifikasi Risiko Kejadian Berat Badan Lahir Rendah di Indonesia. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 20(2), 417–426. <https://doi.org/10.30812/matrik.v20i2.1174>

Taufiq, R., Wardoyo, A. E., & Pratama, R. (t.t.). *ANALISIS SENTIMEN PADA TWITTER TERHADAP KINERJA KOMISI PEMBERANTASAN KORUPSI (KPK) DI INDONESIA DENGAN METODE NAIVE BAYES*.

Utomo, D. P., & Mesran, M. (2020). Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 4(2), 437. <https://doi.org/10.30865/mib.v4i2.2080>

